

What is claimed is

1. A method of covalently joining a DNA strand to an RNA strand comprising:

5 (a) forming a topoisomerase-DNA intermediate by incubating a DNA cleavage substrate comprising a topoisomerase cleavage site with a topoisomerase specific for that site, wherein the topoisomerase-DNA intermediate has one or more 5' single-strand tails; and

10 (b) adding to the topoisomerase-DNA intermediate an acceptor RNA strand complementary to the 5' single-strand tail under conditions permitting a ligation of the covalently bound DNA strand of the topoisomerase-DNA intermediate to the RNA  
15 acceptor strand and dissociation of the topoisomerase, thereby covalently joining the DNA strand to the RNA strand.

2. A method of claim 1, wherein the DNA cleavage substrate is created by hybridizing a DNA strand  
20 having a topoisomerase cleavage site to a complementary DNA strand, thereby forming a DNA cleavage substrate having a topoisomerase cleavage site and a oligonucleotide leaving group located 3' of a scissile bond.

- 25 3. A method of claim 1, wherein the DNA cleavage substrate is a plasmid vector comprising a topoisomerase cleavage site.

4. The method of claim 1, wherein the topoisomerase

cleavage site is a sequence comprising CCCTT.

5. The method of claim 1, wherein the topoisomerase is a vaccinia topoisomerase enzyme.
- 5 6. The method of claim 1, wherein the DNA strand comprising a topoisomerase cleavage site is radiolabelled.
7. The method of claim 6, wherein the radiolabel is  $^{32}\text{P}$  or a radiohalogen.
- 10 8. The method of claim 1, wherein the DNA strand having a topoisomerase cleavage site is labeled with a biotin moiety.
9. The method of claim 1, wherein the topoisomerase-bound DNA intermediate and the acceptor RNA strand are ligated in vitro.
- 15 10. A topoisomerase-DNA intermediate molecule comprising one or more 5' single-strand tails.
11. The topoisomerase-DNA intermediate molecule of claim 10, wherein the 5' single-strand tail comprises a specific sequence.
- 20 12. A topoisomerase-DNA intermediate molecule comprising a 5' single-strand tail generated by step (a) of the method of claim 1.
13. A topoisomerase-DNA intermediate molecule comprising

a 5' single-strand tail generated by steps (a) of the method of claim 1, wherein the 5' single-strand tail comprises a specific sequence.

- 5 14. A topoisomerase-DNA intermediate molecule comprising a 5' single-strand tail generated by steps (a) of the method of claim 1, wherein the DNA strand is radiolabelled.
15. The topoisomerase-DNA intermediate molecule of claim 13, wherein the radiolabel is  $^{32}\text{P}$  or a radiohalogen.
- 10 16. A topoisomerase-DNA intermediate molecule comprising a 5' single-strand tail generated by steps (a) of the method of claim 1, wherein the DNA strand is affinity labeled.
- 15 17. The topoisomerase-DNA intermediate molecule of claim 16, wherein the affinity label is a biotin moiety, a chitin binding domain or a glutathione-S-transferase moiety.
18. A DNA-RNA molecule covalently joined by topoisomerase catalysis.
- 20 19. A DNA-RNA molecule covalently joined by the method of claim 1.
20. The covalently joined DNA-RNA molecule of claim 19, having a 5' end label.
- 25 21. The covalently joined DNA-RNA molecule of claim 20, wherein the 5' end label is  $^{32}\text{P}$  or a radiohalogen.

22. The covalently joined DNA-RNA molecule of claim 20, wherein the 5' end label is a biotin moiety, a chitin binding domain, or a glutathione-S-transferase moiety.
- 5 23. A covalently joined DNA-RNA molecule having a labeled 5' end.
24. The covalently joined DNA-RNA molecule of claim 23, wherein the 5' end label is  $^{32}\text{P}$  or a radiohalogen.
- 10 25. The covalently joined DNA-RNA molecule of claim 23, wherein the 5' end label is a biotin moiety, a chitin binding domain, or a glutathione-S-transferase moiety.
26. A method of tagging a 5' end of an RNA molecule comprising:
- 15 (a) forming a topoisomerase-DNA intermediate by incubating a DNA cleavage substrate comprising a topoisomerase cleavage site with a topoisomerase specific for that site, wherein the topoisomerase-DNA intermediate has one or more 5' single-strand tails; and
- 20 (b) adding to the topoisomerase-DNA intermediate a 5'-hydroxyl terminated RNA molecule complementary to the 5' single-strand tail under conditions permitting a ligation of the covalently bound DNA strand of the topoisomerase-DNA intermediate to the RNA molecule and dissociation of the
- 25 topoisomerase, thereby forming a 5' end tagged DNA-RNA ligation product.
27. A method of claim 26, wherein the 5'-hydroxyl

terminated RNA molecule is the product of in vitro synthesis or isolation from cells or tissues.

28. The method of claim 27, wherein the RNA molecule is dephosphorylated after synthesis or isolation.

5 29. The method of claim 28, wherein the dephosphorylation is achieved by treatment of the RNA molecule with alkaline phosphatase.

10 30. A method of claim 26, wherein the DNA cleavage substrate is created by hybridizing a DNA strand having a topoisomerase cleavage site to a complementary DNA strand, thereby forming a DNA cleavage substrate having a topoisomerase cleavage site and an oligonucleotide leaving group located 3' of a scissile bond.

15 31. The method of claim 26, wherein the topoisomerase is a vaccinia topoisomerase enzyme.

32. The method of claim 26, wherein the cleavage site comprises CCCTT.

20 33. The method of claim 26, wherein the DNA comprises a 5' end label.

34. The method of claim 33, wherein the 5' end label is a biotin moiety, a chitin binding domain, or a glutathione-S-transferase moiety.

25 35. The method of claim 33, further comprising immobilizing the 5' end labeled DNA on a solid support prior to the addition of the 5'-hydroxyl terminated RNA molecule.

36. The method of claim 35, wherein the solid support comprises streptavidin, avidin, chitin or glutathione.

37. The method of claim 35, further comprising, purifying a biotinylated 5' end tagged DNA-RNA ligation product by separating the solid support to which the 5' end labeled DNA-RNA ligation product is immobilized from a liquid phase comprising unmodified RNA.
38. A 5' end tagged RNA molecule.
39. The 5' end tagged RNA molecule of claim 38, wherein the tag is a DNA sequence.
40. The 5' end tagged RNA molecule of claim 39, further comprising a 5' end label.
41. The 5' end tagged RNA molecule of claim 41, wherein the label is  $^{32}\text{P}$  or a radiohalogen.
42. The 5' end tagged RNA molecule of claim 43, wherein the label is a biotin moiety, a chitin binding domain, or a glutathione-S-transferase moiety.
43. A 5' end tagged RNA molecule generated by the method of claim 26.
44. A DNA-RNA molecule which has been joined in vitro by the use of a topoisomerase.
45. A method of obtaining full-length gene sequences comprising:
- (a) isolating full-length mRNA;
  - (b) attaching a DNA tag sequence to the isolated mRNA; and
  - (c) synthesizing cDNA using the tagged mRNA as a template.
46. A method of claim 45, wherein the mRNA is isolated by employing an affinity purification material.
47. A method of claim 46, wherein the mRNA to be isolated

comprises an affinity purification tagged cap structure.

- 5           48. A method of claim 46, wherein the affinity purification tag is a biotin moiety, a chitin binding domain or a glutathione-S-transferase moiety.
49. A method of claim 46, wherein the affinity purification material comprises a solid support complexed with phenylboronic acid, streptavidin, avidin, chitin or glutathione.
- 10          50. A method of claim 49, wherein the solid support is magnetic beads or sepharose.
51. A method of claim 45 wherein the mRNA is isolated from plant cells or animal cells.
- 15          52. A method of claim 51 wherein the animal cells are mammalian cells or insect cells.
53. A method of claim 45, wherein the mRNA is decapped and dephosphorylated after isolation.
54. A method of claim 53 wherein the mRNA is decapped enzymatically or by chemical treatment.
- 20          55. A method of claim 54 wherein the enzyme is a pyrophosphatase.
56. A method of claim 54 wherein the chemical treatment is periodate oxidation and beta elimination.
- 25          57. A method of claim 53 wherein the mRNA is dephosphorylated using alkaline phosphatase.
58. A method of claim 45, wherein the DNA tag sequence comprises a recognition site for a type I topoisomerase.

59. A method of claim 58 wherein the DNA tag sequence further comprises a recognition site for a site-specific restriction endonuclease.
- 5 60. A method of claim 58 wherein the type I topoisomerase is vaccinia DNA topoisomerase.
61. A method of claim 58 wherein the DNA tag sequence comprises the double stranded sequence shown in Figure 11 wherein N represents an adenosine moiety, a guanosine moiety, a cytosine moiety or a thymidine moiety.
- 10 62. A method of claim 61 wherein N is 1 to 4 nucleotide bases.
63. A method of claim 61 wherein vaccinia DNA topoisomerase is covalently bound to the double stranded tag sequence.
- 15 64. A method of claim 45 further comprising amplifying the synthesized cDNA wherein the amplification primers comprise an anti-coding sequence of the tag sequence (5') and a gene specific sequence (3').
- 20 65. A method of claim 64 further comprising inserting the amplified cDNA into an expression vector.
66. A method of claim 45 wherein the DNA tag sequence is a linearized expression vector.
- 25 67. An isolated full-length gene sequence prepared by the method of claim 45.
68. A nucleic acid construct comprising an isolated full-length gene sequence prepared of the method of claim 45 and an expression vector.
69. A nucleic acid construct of claim 68 wherein the

expression vector comprises one or more elements selected from: a promoter-enhancer sequence, a selection marker sequence, an origin of replication, an epitope-tag encoding sequence or an affinity purification-tag encoding sequence.

5

70. A nucleic acid construct of claim 69 wherein the promoter-enhancer sequence is the T7 promoter, gall promoter, metallothionein promoter, AraC promoter, or CMV promoter-enhancer.

10

71. A nucleic acid construct of claim 69 wherein the selection marker sequence encodes an antibiotic resistance gene.

72. A nucleic acid construct of claim 69 wherein the epitope-tag sequence encodes V5, the peptide Phe-His-His-Thr-Thr, hemagglutinin, or glutathione-S-transferase.

15

73. A nucleic acid construct of claim 69 wherein the affinity purification-tag sequence encodes a polyamino acid sequence or a polypeptide.

20

74. A nucleic acid construct of claim 73 wherein said polyamino acid sequence is polyhistidine.

75. A nucleic acid construct of claim 73 wherein said polypeptide is chitin binding domain or glutathione-S-transferase.

25

76. A nucleic acid construct of claim 73 wherein said polypeptide encoding sequence includes an intein encoding sequence.

30

77. A nucleic acid construct of claim 68 wherein the expression vector is a eukaryotic expression vector or a prokaryotic expression vector.

78. A nucleic acid construct of claim 77 wherein the eukaryotic expression vector is pYES2, pMT, pIND, or pcDNA3.1.

5 79. A method of obtaining full-length gene sequences comprising:

(a) isolating full-length mRNA by employing an affinity purification material;

(b) decapping and dephosphorylating the isolated mRNA;

10 (c) attaching a DNA tag sequence to the decapped, dephosphorylated mRNA, wherein the tag sequence comprises the sequence shown in Figure 11 and is attached by vaccinia DNA topoisomerase;

15 (d) synthesizing cDNA using the tagged mRNA as a template;

(e) amplifying the synthesized cDNA, wherein the amplification primers comprise an anti-coding sequence of the tag sequence (5') and a gene specific sequence (3'); and

20 (f) inserting the amplified cDNA into an expression vector.